PRS: Physics Reconstruction and Selection
HCAL/JetsMET group

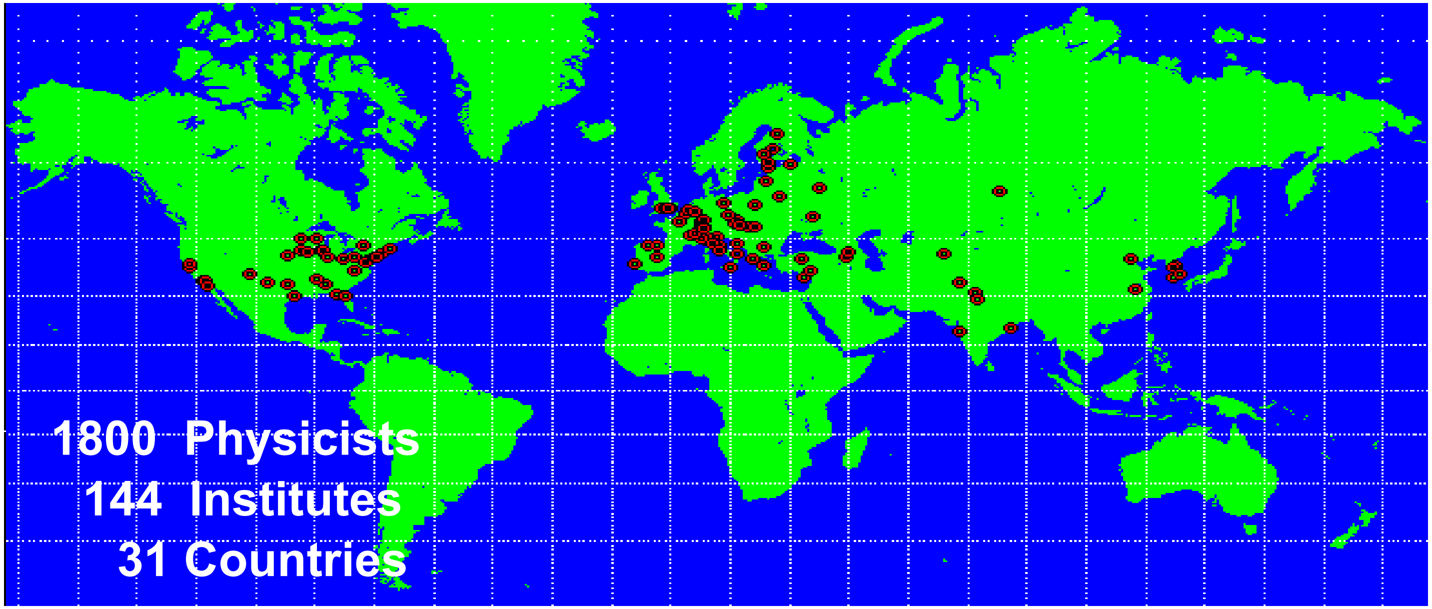
CMS Computing/Software

Shuichi Kunori
U. of Maryland

19-Feb-2002

LHC Computing Challenges

- **Geographical Dispersion:** of people and resources
- **Complexity:** the detector and the LHC environment
- **Scale:** Petabytes per year of data

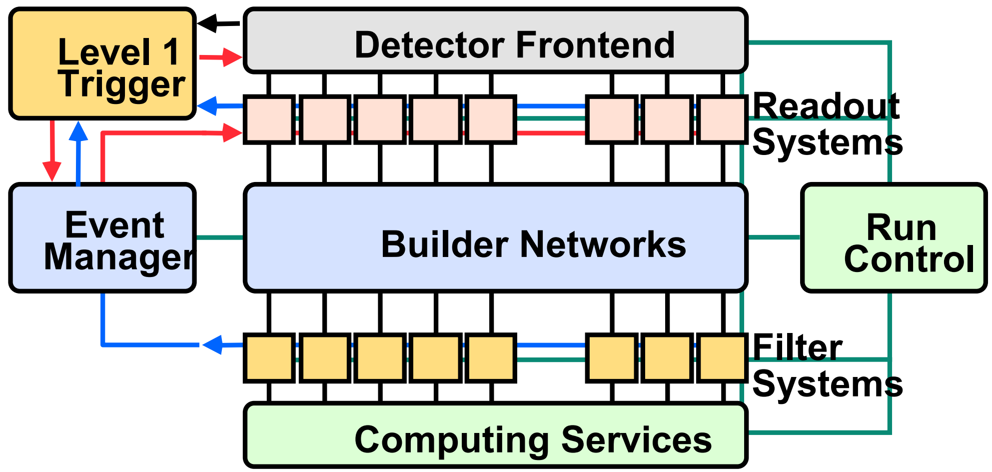


Major challenges associated with:
Communication and collaboration at a distance
Network-distributed computing and data resources
Remote software development and physics analysis
R&D: New Forms of Distributed Systems: Data Grids

CMS Trigger and Data Acquisition

40 MHz

100 Tbyte/s



~~10^5 Hz~~

L1 output:75kHz

100 Gbyte/s

~~10^2 Hz~~

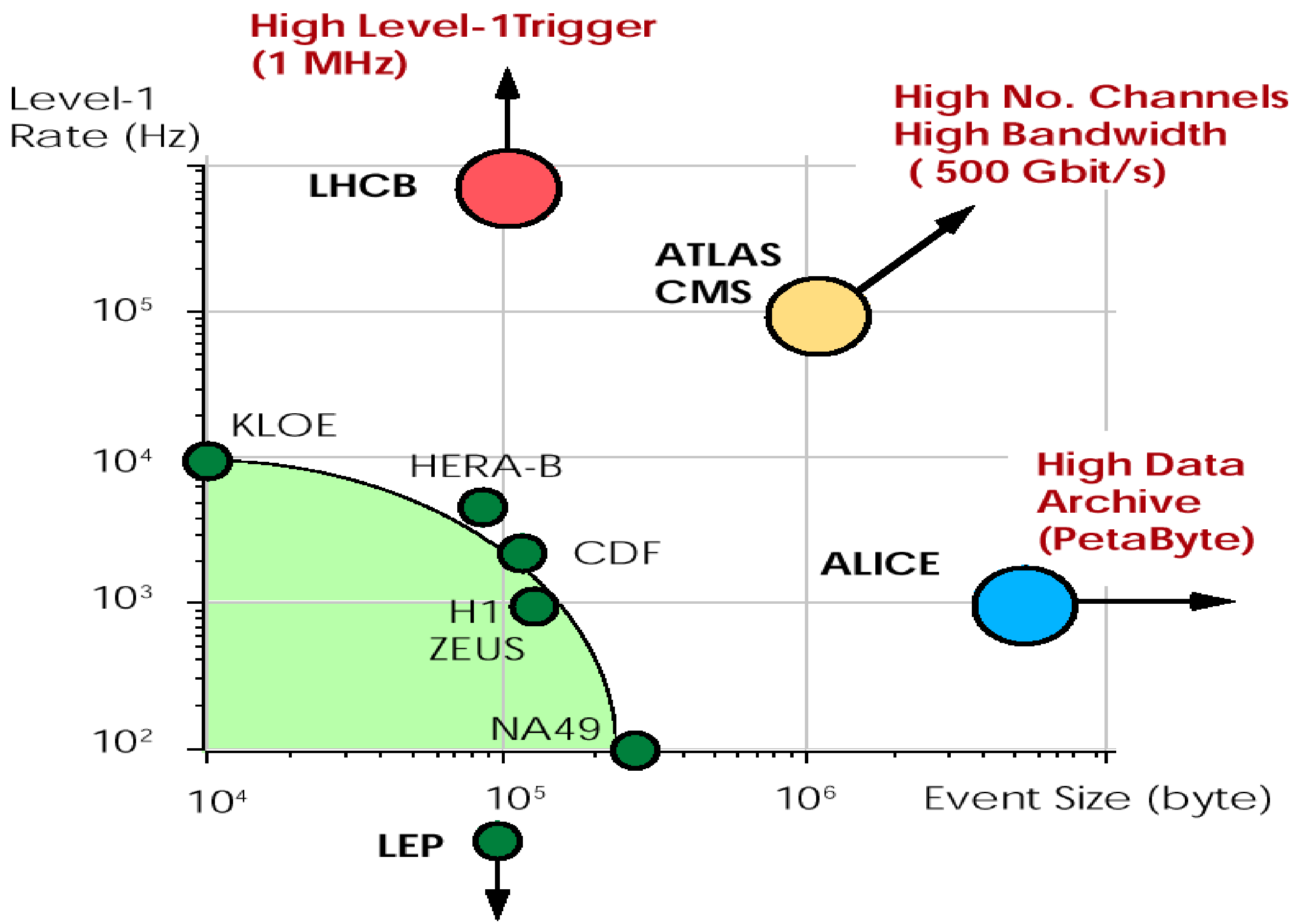
HLT output: 75Hz

100 Mbyte/s

Collision rate	40 MHz
Level-1 Maximum trigger rate	100 kHz
Average event size	1 Mbyte
Event Flow Control	~10^6 Msg/sec
No. of In-Out units (200-5000 byte/event)	1000
Readout network (512-512 switch) BW	1 Terabit/s
Event filter computing power	~200 kSI95
Data production	~10 Tbyte/day
No. of readout crates	250
No. of electronics boards	10000

Data Rate and Size

Larger by ~2 Orders of Magnitude!



Distributed Computing
(CMS Analysis Model)

"Reconstruction" (first pass) creates ESD (event summary data) from the RAW data

- Once at CERN, synchronized with data taking

"Re-processing" re-creates ESD data from the RAW data and from the "old" ESD

- 3 times per year, about 3 months response time per pass

Scheduled

"Selection" creates AOD (analysis object data) and TAG from the ESD data

- Once per month per Analysis Group

"Analysis"

- Creates DPD (Derived Physics Data), personal dataset
- Iterate over the DPD data and over the TAG/AOD data
- Eventually access a fraction of the ESD data
  (it's also possible to permit, under control, access a fraction of the RAW data)

Fully distributed

Chaotic

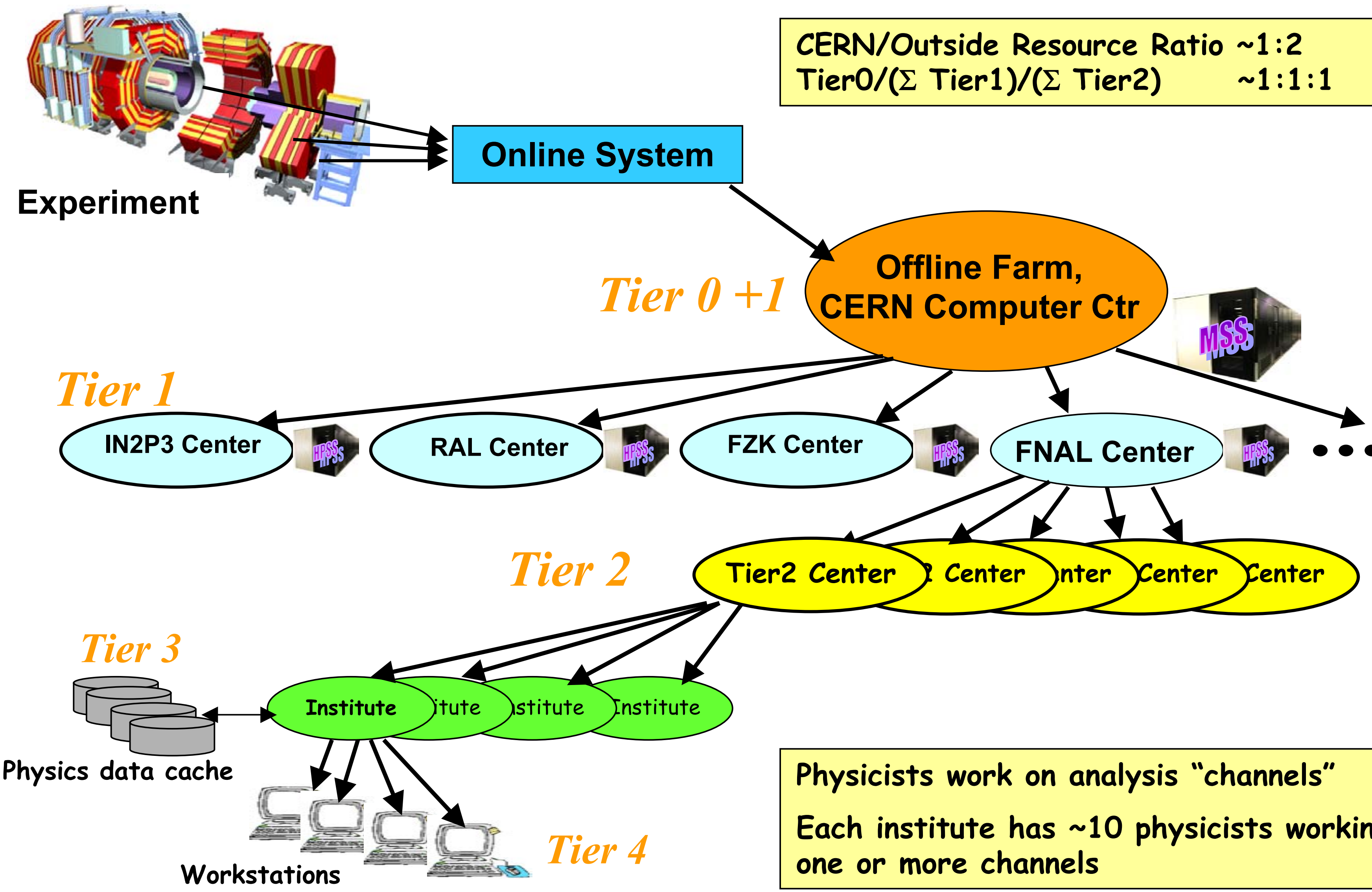
"Monte Carlo Simulation" creates RAW-like data (MCRAW) (+ Pile-up simulation)

- Same process as for real RAW data: ESD, AOD etc.!
- A total of about 5×10^8 events per year

Scheduled

CMS Computing Model: Data Grid Hierarchy

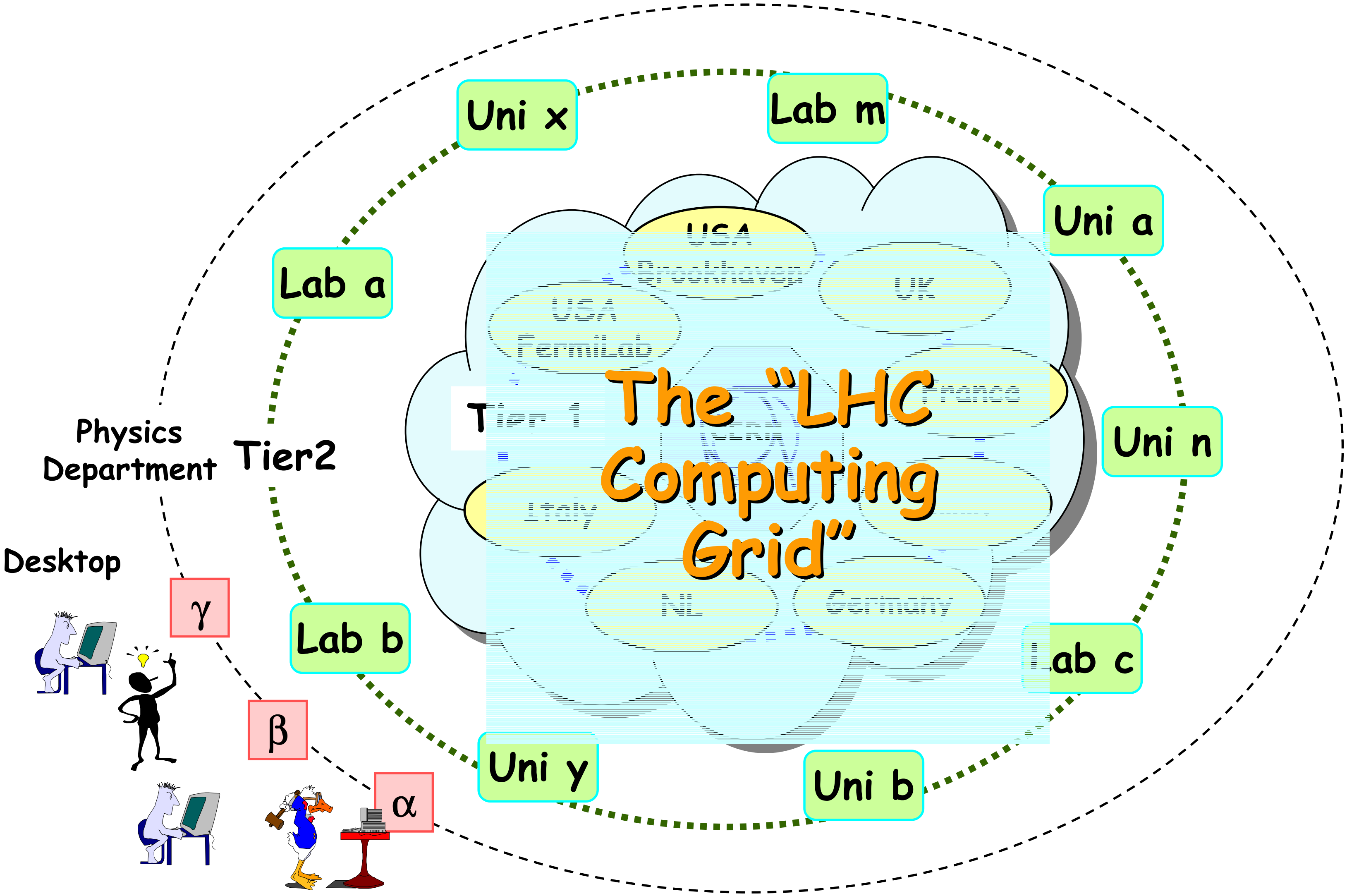
CERN/Outside Resource Ratio ~1:2
Tier0/(Σ Tier1)/(Σ Tier2) ~1:1:1



Physicists work on analysis "channels"

Each institute has ~10 physicists working on one or more channels

LHC Distributed Computing Model



A few of the Grid Technology Projects

Data-intensive projects with EU funding, HEP leadership

- DataGrid – 21 partners, coordinated by CERN (Fabrizio Gagliardi)
- CrossGrid – 23 partners complementary to DataGrid (Michal Turala)
- DataTAG – funding for transatlantic demonstration Grids (Olivier Martin)

European national HEP related projects

- GridPP (UK); INFN Grid; Dutch Grid; NorduGrid; Hungarian Grid; ……

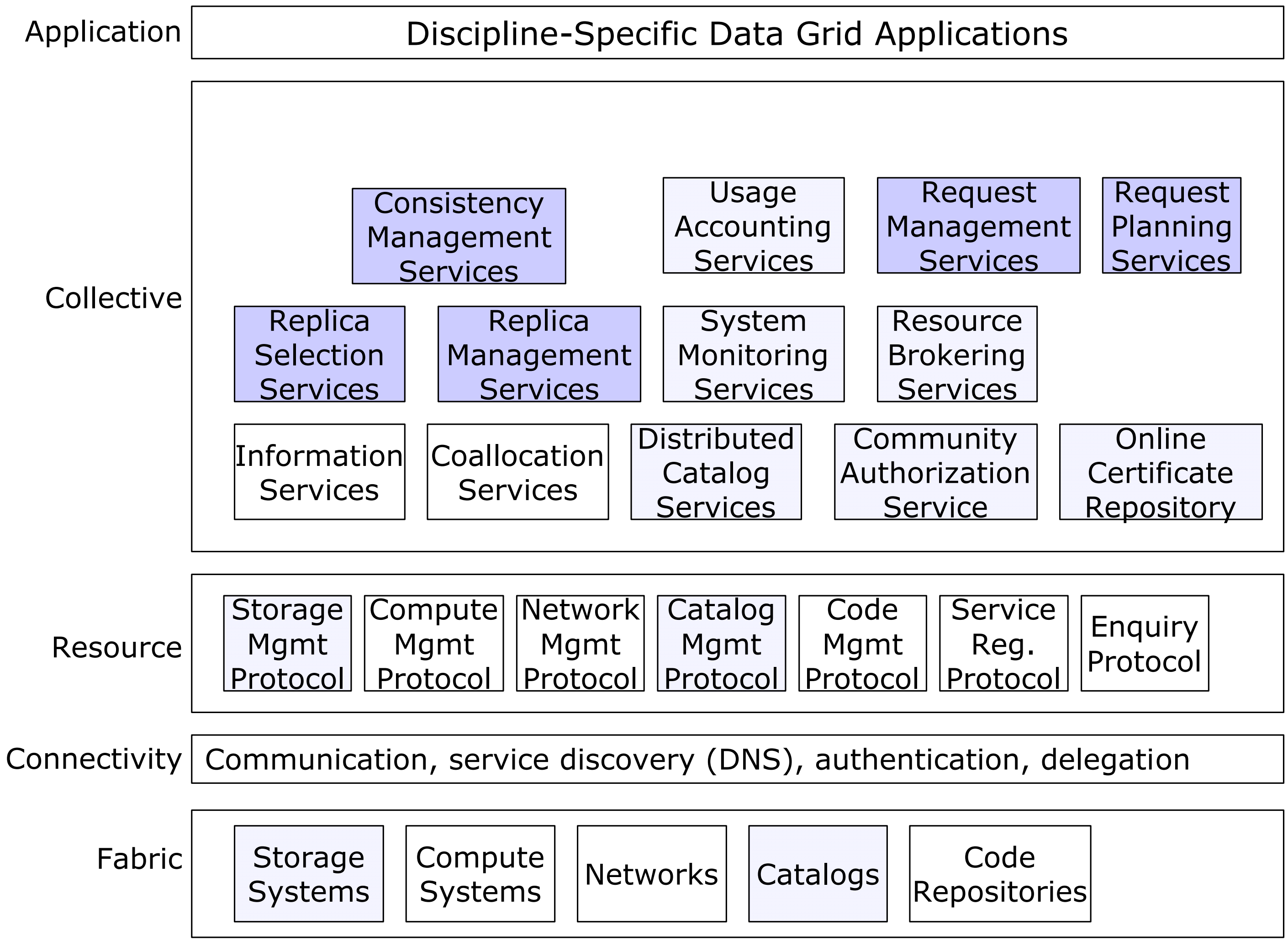
US HEP projects

- GriPhyN – NSF funding; HEP applications
- PPDG – Particle Physics Data Grid – DoE funding
- iVDGL – international Virtual Data Grid Laboratory

Global Coordination

- Global Grid Forum – GGF Recommendations
- InterGrid – ad hoc HENP Grid coordination (Larry Price)

Data Grid Reference Architecture



'Data Grid' is used to describe system with access to large volumes of data

Grid Services for the User

What does the Grid do for you?

- **you submit your work**
- **and the Grid**
  - Finds convenient places for it to be run
  - Optimises use of the widely dispersed resources
  - Organises efficient access to your data
    - **Caching, migration, replication**
  - Deals with authentication to the different sites that you will be using
  - Interfaces to local site resource allocation mechanisms, policies
  - Runs your jobs
  - Monitors progress
  - Recovers from problems

  - .. and .. Tells you when your work is complete

Short Term CMS/GRID Issues

CMS input to GRID projects

- CMS note describing our initial requirements on GRID architecture (CMS NOTE-2001/037)

GDMP (Grid Data Managgement Pilot) as a practical GRID tool

- Built by Computer Science Students in CMS
  - Initial version, was limited to transferring Objectivity database files. More recently significantly extended GDMP capabilities by integrating two new Globus Data Grid tools (Globus Replica Catalog, GridFTP)
- Widely deployed in CMS productions.
  - CERN, INFN, FNAL, UK, UCSD, Caltech, Moscow,…

Obtaining coherence in worldwide GRID projects

- CMS is worldwide, not American, not European, not Asian
  - We need coherent middleware, or a coherent interface layer to the various middleware

Production Sites

	Simulation	Digitization		GDMP	Common Production tools (IMPALA)
		No PU	PU		
CERN				✓	✓
FNAL				✓	✓
Moscow				✓	In progress
INFN	Fully operational			✓	✓
Caltech				✓	✓
UCSD				✓	✓
UFL				✓	✓
Imperial College				✓	✓
Bristol				✓	✓
Wisconsin				✓	✓
IN2P3			Not Op.	✓	✓
Helsinki		Not Op.	Not Op.		

Prototyping and Production today

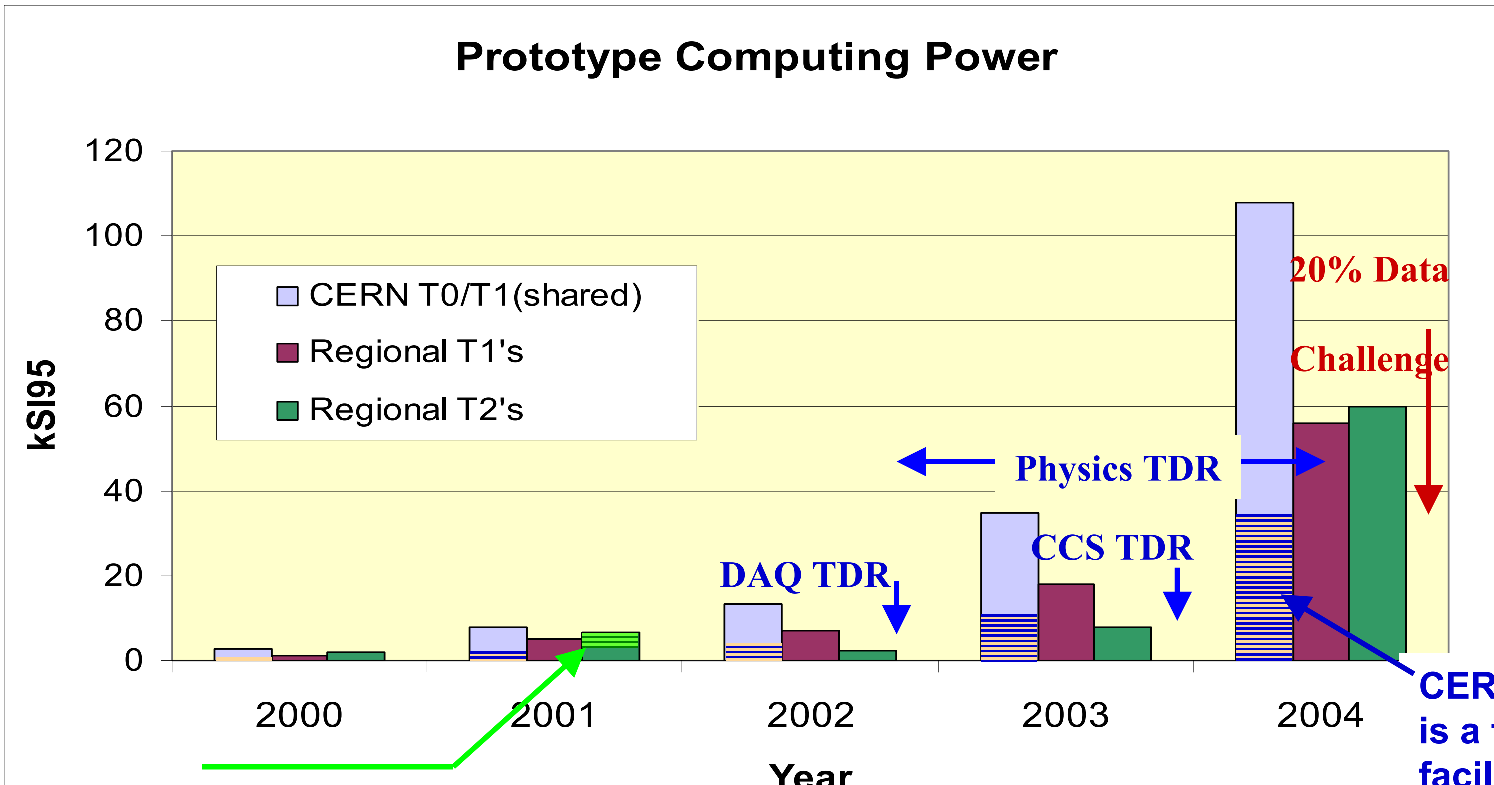
- We need large scale computing to:
  - ➔Satisfy computing requirements for DAQ TDR and Physics TDR
  - ➔Test Software components under realistic conditions
  - ➔Develop ways to efficiently use the computing available
- Target to reach 50% of complexity by 2004
  - ➔T0/T1 with approx 600 “cpu boxes” at CERN
  - ➔ Likewise scale offsite T1 prototypes (assume 2) and offsite T2 prototypes (assume 10)
  - ➔ 20% Data Challenge in 2004
- Data Challenges require worldwide resources operating cooperatively
  - ➔Efficient operation of ~100’s of Tiers3’s in this environment is untested

Lesson from the Prototypes

- (Economic, large scale) Data access is the challenge
  - ➔High disk failure rates
  - ➔Mass-Storage software not yet mature
  - ➔Complex systems built from many “cheapest” components
  - ➔Problems for Productions and for Analysis
- The data challenges and productions are R&D.
  - ➔We do not get things right first time
  - ➔Both CMS and the computing centers learn from these exercises
  - ➔Build and share experiences in different environments (T0->T3)

Common Prototypes: CMS Computing, 2002-2004

Match Computing Challenges with CMS Physics and Detector Milestones



Some (~50%) of current T2 prototypes primarily for GRID related R&D.

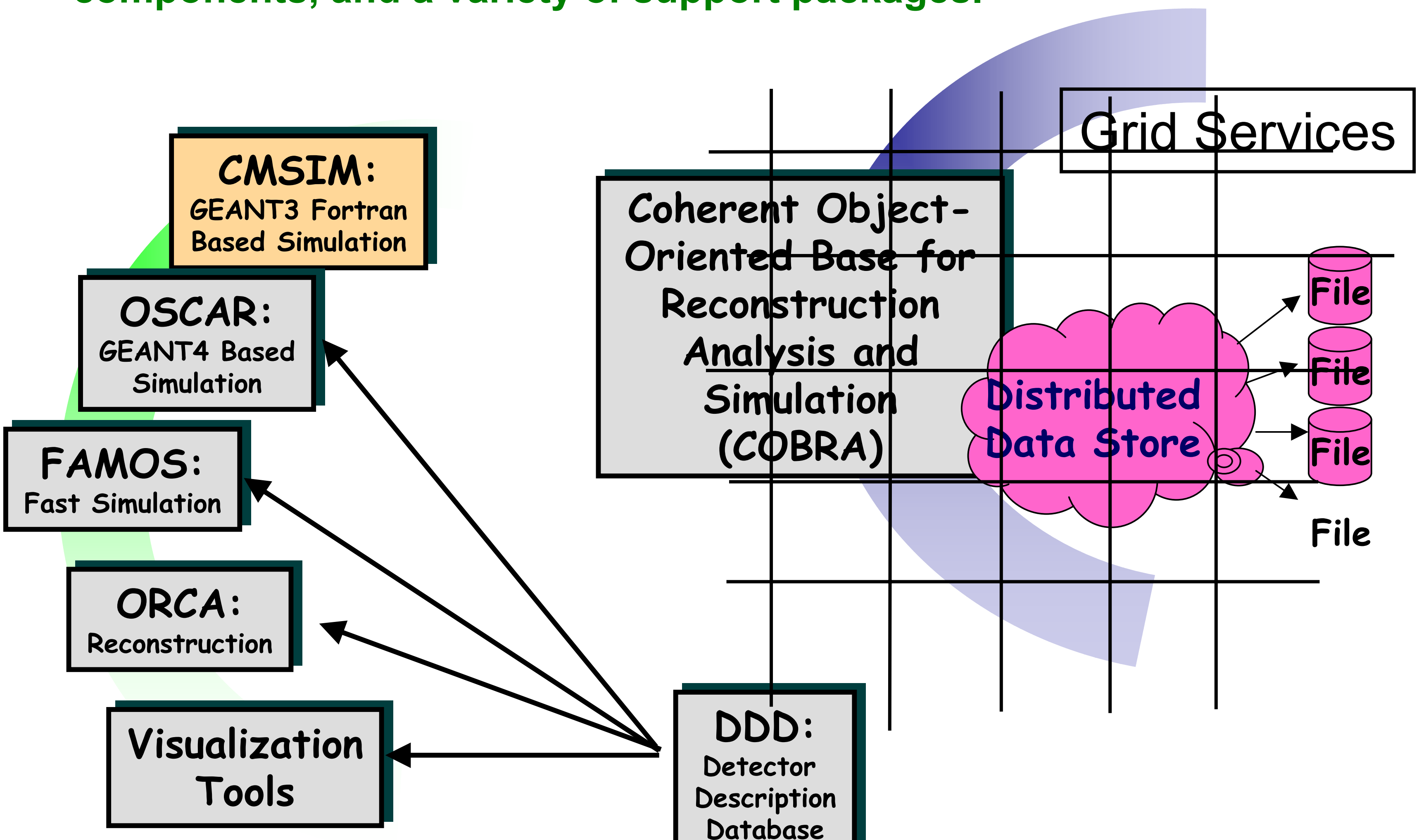
CERN prototype is a time-shared facility available for ~30% of the time at full power for CMS

Software Project

- CMSIM
  - → The original GEANT3 Simulation of CMS.
- COBRA (Coherent Object-oriented Base for simulation, Reconstruction and Analysis)
- ORCA (Object Reconstruction for CMS Analysis)
  - → The OO reconstruction program
- OSCAR (Object oriented Simulation for CMS Analysis and Reconstruction)
  - → The GEANT4 Simulation framework for CMS
- IGUANA (Interactive Graphical User ANAlysis)
  - → Toolkits for Interactive Analysis
- FAMOS (Fast Monte-Carlo Simulation)
  - → "Parameterized" Monte-Carlo
- DDD (needs an Acronym!)
  - → The Detector Description Database
- GDMP, MOP, CLARENS, BOSS,…
  - → Grid projects with strong CMS involvement/authorship
- IMPALA
  - → Production Tools

Software Components

CMS Software has a data store, a central framework, a number of components, and a variety of support packages.



Objectivity

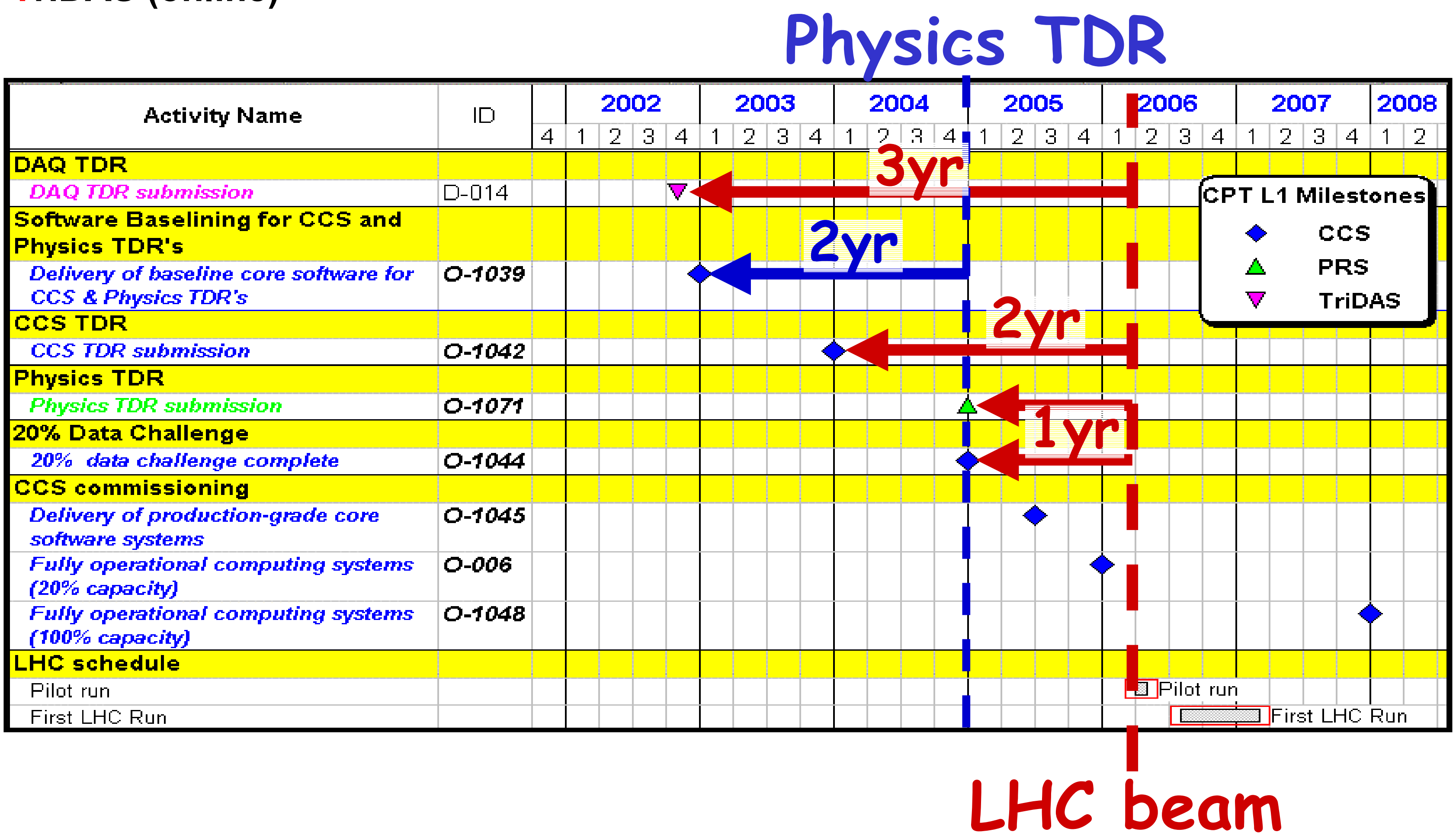
- **Objectivity**
  - ➔Currently about 30TB in Objectivity DB's
  - ➔Experience with writing into DB with up to 300 CPU's in parallel
  - ➔Little experience to date with large numbers of parallel readers
  - ➔We have confidence that we *could* make an Objectivity based solution work
- **Commercial Considerations**
  - ➔Object databases have not taken off as forecast
  - ➔Objectivity is the only major vendor of an ODBMS
  - ➔Difficult to support long term in the event the company fails
- **Observations**
  - ➔(backed up by discussions with Babar, and initial CMS/Oracle studies)
    - The user code, is not the big issue; we can hide almost all of this from the users
    - However the data management is optimized for the specific implementation and this is where a changeover will be most painful

ROOT IO + DB

- http://cmsdoc.cern.ch/cms/ccs/bulletin/bulletin-dec-2001.html
- Document Objectivity experience. Develop evaluation criteria
- Concentrate current work on Hybrid scheme (ROOT+DB Layer)
  - ➔Develop overall plan
  - ➔Evaluate solution
  - ➔Estimate work required
  - ➔Identify partners
- Establish a working group together with ROOT, IT/DB, and all LHC experiments.
- Specifying a "HEPIO" format based on "ROOTIO"
  - Establish a change control mechanism
  - Allow for other products to be developed against the specification
  - (FNAL taking responsibility for this)
- *Adopt* rather than *adapt*.

CPT Level 1 Milestones

Core Software and Computing
Physics Reconstruction and Selection
TriDAS (online)



CCS Organization

- Management
  - Project Manager:
    - Martti Pimiä (CERN) ,
    - David Stickland (Princeton), from Jan 1 2002
  - Resource Manager. Ian Willers (CERN)
  - Technical Coordinator. Lucas Taylor (Northeastern)
  - CCS-IB Chair. Harvey Newman (Caltech)

- Level-2 tasks
  - Computing:
    - Martti Pimiä (CERN)
  - General Services:
    - Werner Jank (CERN)
  - Architecture, Frameworks and Toolkits:
    - Vincenzo Innocente (CERN)
  - Developers and Users Environment
    - Stephan Wynhoff (Princeton) (Ad interim)
  - Software Process
    - Johannes-Peter Wellisch (CERN)
  - Productions and Data Management
    - Tony Wildish (Princeton)
  - GRID Integration
    - Claudio Grandi (INFN Bologna)

- LHC Computing Project
  - Project Oversight Board:
    - Michel Della Negra (CERN)
  - Project Execution Board:
    - Lucas Taylor (Northeastern)
  - Software and Computing Committee (SC2)
    - David Stickland (Princeton)
    - Paolo Capiluppi (INFN/Bologna)

CPT Organization

- Three CMS projects working together:
  - → CCS: Core Software and Computing
  - → PRS: Physics Reconstruction an Selection
  - → TriDAS(Online): Online Computing and Farm related tasks of DAQ
- Each has a project manager in the CMS Steering Committee
- Regular inter-PM meetings to ensure coherence.
- One Joint Technical Board
  - → L1 and L2 of CCS, PRS, TriDAS(Online)
- Cross-Project Task Forces as required
  - → Reconstruction:
    - ◻ Stephan Wynhoff (Princeton)
  - → Simulation:
    - ◻ Albert de Roeck (CERN)
  - → SW Process Improvement
    - ◻ Johannes-Peter Wellisch (CERN)